(Puregene® DNA isolation kit for yeast and Gram positive bacteria, Gentra systems Minneapolis, MN). The forward primer (5' ATGGATTATTATAGAGATGTGAG 3'), was based on sequence from the genome sequencing project, starting 112 nucleotides upstream from the start of the *slpA* open reading frame. Two reverse primers were used, depending on the PCR type. A downstream primer (5' CTATTTAAAGTTTTATTAAAACTTATATTAC 3') was used to amplify *slpA* from PCR types 12, 17, 31, 46 and 92. A reverse primer based on the 3' end of the *slpA* open reading frame from strain 630 and the subsequent nonsense codon (5' TTACATATCTAATAAATCTTTCATTTTGTTTATAACTG 3') was used to amplify *slpA* from PCR types 1 and 5. The choice of primer for the latter two PCR types may have resulted in a small number of systematic errors in the nucleotide sequence obtained. PCR was carried out using HotStar™ Taq polymerase (Qiagen Ltd., Crawley, West Sussex, UK) according to the manufacturer's instructions. A single fragment of approximately 2 kb was obtained for each strain, which was then cloned into the pBAD/Thio TOPO vector (Invitrogen, Groningen, Netherlands). Inserts were sequenced from both ends by standard procedures in commercial facilities at MWG (Wolverton Mill South, Milton Keynes, UK) and Cambridge University. New primers were designed on the basis of initial sequencing results, enabling sequencing of both strands to be completed (a process known as chromosome walking).

The results are shown in Appendices 1-8.

The nucleotide sequences were translated to enable prediction of the amino acid sequence(s) of the product(s) (Appendices 1-8). The N-terminal sequences obtained experimentally for the low molecular weight protective antigens from strains 171500 (PCR type 1) and 170324 (PCR type 12) were almost identical to those predicted from the nucleotide sequences of their respective *slpA* genes (18/20 identical residues for strain 171500, and 19/20 identical residues for strain 170324).

Appendix 1 shows the open reading frame with translation for *slpA* from strain 171500 (PCR type 1), SEQ ID No 3. Since the reverse primer was based on the 35 nucleotides from the 3' end of the *slpA* gene, the sequence is not necessarily 100% accurate in this region. However, this part of the gene does not seem to vary greatly from strain to strain.